

Об итеративном алгоритме подсчета взвешенных главных компонент

Е.В. Бурнаев*, С.С. Чернова**

*Институт проблем передач информации им. А.А. Харкевича РАН, Москва, Россия

**Институт системного анализа РАН, Москва, Россия

Поступила в редколлегию 06.05.2008

Аннотация—Разработан итеративный алгоритм подсчета взвешенных главных компонент, который позволяет существенно уменьшить среднюю (по выборке) взвешенную максимальную (по координате) ошибку восстановления векторов выборки по сжатым данным. Приведены результаты моделирования, иллюстрирующие эффективность алгоритма.

1. ВВЕДЕНИЕ

Снижение размерности данных представляет собой такое преобразование многомерных данных, которое снижает размерность данных до внутренней размерности или близкой к ней (под внутренней размерностью данных понимается такое количество параметров, которое достаточно для того, чтобы объяснить наблюдаемые свойства данных [1, 2, 3, 4]).

Снижение размерности широко применяется во многих областях науки и техники, поскольку оно позволяет существенно увеличить эффективность различных алгоритмов обработки данных за счет ослабления нежелательных эффектов, вызванных “проклятием размерности” [2, 5, 6, 7, 8].

Метод главных компонент (ГК), по всей видимости, в силу своей простоты, аналитических свойств и наличия эффективных алгоритмов подсчета является наиболее распространенным методом снижения размерности [9, 10, 11].

Существует достаточно большое количество модификаций ГК, учитывающих те или иные практические нужды. Например, в [12, 13] рассмотрены подходы к подсчету ГК для данных, принимающих дискретные значения. В [14, 15] предложены варианты метода ГК, с помощью которых можно обрабатывать неполные данные. В [16, 17] разработаны алгоритмы подсчета ГК, которые позволяют получать разреженную матрицу нагрузок (матрица, с помощью которой вычисляются ГК).

Известно, что ГК минимизируют сумму квадратов евклидовых расстояний между исходными векторами выборки и их восстановленными аналогами [11]. Однако, классические алгоритмы подсчета ГК подвержены влиянию выбросов, поэтому в [14, 18, 19] были предложены робастные варианты подсчета ГК. Также в [14, 18, 19] за счет введения весов предлагалось: учитывать “надежность” того или иного наблюдения из выборки, используемой для оценки ГК, а также влиять на точность восстановления определенных координат по сжатым векторам.

Однако, до сих пор ни в одной из работ не был разработан алгоритм подсчета ГК, который бы позволял минимизировать среднюю (по выборке) взвешенную максимальную (по координате) ошибку восстановления векторов выборки по сжатым данным. В данной статье предлагается один из возможных подходов к построению такого алгоритма.

В статье рассматривается постановка задачи снижения размерности данных (раздел 2), приводится математическая постановка решаемой задачи (раздел 3), описывается алгоритм,

решающий поставленную задачу (раздел 4), на реальных данных иллюстрируется применение предложенного алгоритма (раздел 5), обсуждаются полученные результаты и предлагаются пути дальнейших исследований (раздел 6).

2. СНИЖЕНИЕ РАЗМЕРНОСТИ ДАННЫХ

Имеется выборка $\{\mathbf{X}_t\}_{t=1}^T$ N -мерных вектор-столбцов, лежащих в пространстве данных \mathcal{X} (обычно это пространство \mathbb{R}^N или его подмножество). В дальнейшем для простоты будем считать, что выборка центрирована, т.е. $\sum_{t=1}^T \mathbf{X}_t = \mathbf{0} \in \mathbb{R}^N$. Фундаментальное предположение, которое обосновывает возможность снижения размерности, состоит в том, что выборка в действительности лежит, по крайней мере приближенно, на многообразии (в общем случае нелинейном), размерность которого меньше, чем размерность пространства данных. Задача снижения размерности заключается в том, чтобы построить представление этого многообразия (его координатную систему) с учетом дополнительных требований, предъявляемых к процедурам восстановления.

Формально, снижение размерности состоит в следующем. Для данной выборки $\{\mathbf{X}_t\}_{t=1}^T \subset \mathcal{X}$ необходимо найти

1. Пространство \mathcal{Y} размерности n (обычно это \mathbb{R}^n или его подмножество).
2. Снижающее размерность отображение \mathbf{F} ,

$$\begin{aligned} \mathbf{F} : \mathcal{X} &\rightarrow \mathcal{Y} \\ \mathbf{X} &\mapsto \mathbf{Y} = \mathbf{F}(\mathbf{X}), \end{aligned}$$

результат $\mathbf{Y} = \mathbf{F}(\mathbf{X})$ применения которого к вектору $\mathbf{X} \in \mathcal{X}$ будем называть представлением исходного вектора \mathbf{X} в виде сжатого вектора меньшей размерности.

3. Гладкое, несингулярное¹ восстанавливающее отображение \mathbf{F}^- ,

$$\begin{aligned} \mathbf{F}^- : \mathcal{Y} &\rightarrow \mathcal{M} \subset \mathcal{X} \\ \mathbf{Y} &\mapsto \mathbf{X} = \mathbf{F}^-(\mathbf{Y}). \end{aligned}$$

При этом необходимо, чтобы

1. Размерность $n < N$ была как можно меньше.
2. Выполнялось условие, что

$$\text{многообразие } \mathcal{M} \stackrel{\text{def}}{=} \mathbf{F}^-(\mathcal{Y}) \text{ приблизительно содержит выборку.} \quad (1)$$

Это условие можно переформулировать и иначе, а именно:

$$\text{Ошибка восстановления выборки должна быть мала.} \quad (2)$$

Ошибку восстановления выборки определим как

$$\delta_1(n, \rho, \mathbf{w}) = \frac{1}{T} \sum_{t=1}^T \rho(\mathbf{X}_t, \hat{\mathbf{X}}_t, \mathbf{w}), \quad (3)$$

или

$$\delta_2(n, \rho, \mathbf{w}) = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\rho(\mathbf{X}_t, \hat{\mathbf{X}}_t, \mathbf{w}) \right)^2}, \quad (4)$$

¹ Требования гладкости и несингулярности восстанавливающего отображения накладываются с тем, чтобы размерности областей определения и значения восстанавливающего отображения совпадали (несингулярность) и сохранялась топографическая структура области определения (гладкость).

где $\widehat{\mathbf{X}}_t \stackrel{\text{def}}{=} \mathbf{F}^{-1}(\mathbf{F}(\mathbf{X}_t))$ – восстановленный вектор для элемента выборки \mathbf{X}_t , а $\rho(\mathbf{X}_t, \widehat{\mathbf{X}}_t, \mathbf{w})$ – взвешенная с положительными весами $\mathbf{w} = (w_1, \dots, w_N)^\top$ метрика между векторами $\mathbf{X}_t = (X_{1,t}, \dots, X_{N,t})^\top$ и $\widehat{\mathbf{X}}_t = (\widehat{X}_{1,t}, \dots, \widehat{X}_{N,t})^\top$ в пространстве \mathcal{X} . В качестве такой метрики можно использовать, например, следующие метрики:

$$\begin{aligned} \rho_\infty(\mathbf{X}_t, \widehat{\mathbf{X}}_t, \mathbf{w}) &= \max_{j=1, \dots, N} \left\{ w_j \left| X_{j,t} - \widehat{X}_{j,t} \right| \right\}, \\ \rho_1(\mathbf{X}_t, \widehat{\mathbf{X}}_t, \mathbf{w}) &= \frac{1}{N} \sum_{j=1}^N w_j \left| X_{j,t} - \widehat{X}_{j,t} \right|, \\ \rho_2(\mathbf{X}_t, \widehat{\mathbf{X}}_t, \mathbf{w}) &= \sqrt{\frac{1}{N} \sum_{j=1}^N w_j^2 \left(X_{j,t} - \widehat{X}_{j,t} \right)^2}. \end{aligned}$$

Веса \mathbf{w} позволяют учесть, что некоторые координаты векторов выборки $\{\mathbf{X}_t\}_{t=1}^T$ должны быть восстановлены с большей точностью, чем остальные координаты.

Условия (1) и (2) не эквивалентны: из (2) следует (1), но не наоборот. Это происходит потому, что для заданного многообразия $\mathcal{M} = \mathbf{F}^{-1}(\mathcal{Y}) \subset \mathcal{X}$ композиция отображений $\mathbf{F} \circ \mathbf{F}^{-1}$ не обязательно является тождественным отображением. Таким образом, возможна ситуация, что $\{\mathbf{X}_t\}_{t=1}^T \subset \mathcal{M}$, т.е. для каждого \mathbf{X}_t найдется вектор $\mathbf{Y}_t \in \mathcal{Y}$, удовлетворяющий равенству $\mathbf{F}^{-1}(\mathbf{Y}_t) = \mathbf{X}_t$, для которого $\mathbf{F}(\mathbf{X}_t) \neq \mathbf{Y}_t$. Следовательно, $\widehat{\mathbf{X}}_t \stackrel{\text{def}}{=} \mathbf{F}^{-1}(\mathbf{F}(\mathbf{X}_t)) \neq \mathbf{X}_t$ и значение $\delta(n, \rho, \mathbf{w})$ может оказаться большим.

В силу того, что размерности пространств \mathcal{X} и \mathcal{Y} не равны, найдется подмногообразие $(\mathbf{F})^{-1}(\mathbf{Y}) \stackrel{\text{def}}{=} \{\mathbf{X} \in \mathcal{X} : \mathbf{F}(\mathbf{X}) = \mathbf{Y}\}$ размерности $N - n$ в пространстве \mathcal{X} , которое будет отображаться на одну и ту же точку в \mathcal{Y} .

3. ПОСТАНОВКА ЗАДАЧИ

В дальнейшем будем считать, что отображения \mathbf{F} и \mathbf{F}^{-1} – линейные, т.е. $\mathbf{F}(\mathbf{X}) = \mathbf{A}\mathbf{X}$ и $\mathbf{F}^{-1}(\mathbf{Y}) = \mathbf{B}\mathbf{Y}$, где $\mathbf{A} \in \mathbb{R}^{n \times N}$ и $\mathbf{B} \in \mathbb{R}^{N \times n}$.

Пусть $w_j = 1, j = 1, \dots, N$. Тогда решением задачи минимизации ($\mathbf{I} \in \mathbb{R}^{n \times n}$ – единичная матрица)

$$\delta_2(n, \rho_2, \mathbf{w}) \rightarrow \min_{\mathbf{A}, \mathbf{B}: \mathbf{A}\mathbf{B}=\mathbf{I}} \quad (5)$$

являются главные компоненты (строки проекционной матрицы $\mathbf{A} = \mathbf{A}(\mathbf{w})$ равняются компонентам собственных векторов ковариационной матрицы выборки $\{\mathbf{X}_t\}_{t=1}^T$, соответствующих первым n максимальным собственным значениям, при этом $\mathbf{B} = \mathbf{A}^\top$) [9, 10, 11].

Если некоторые $w_j \neq 1$, то решение задачи минимизации (5) получается следующим образом [9, 10, 11]:

1. Пусть элементы матрицы $\mathbf{C}^{\mathbf{w}}$ равны $c_{i,j}^{\mathbf{w}} = c_{i,j} w_i w_j$, где $c_{i,j}$ – элементы ковариационной матрицы \mathbf{C} выборки $\{\mathbf{X}_t\}_{t=1}^T$. Матрица $\mathbf{C}^{\mathbf{w}}$ является ковариационной матрицей выборки $\{\mathbf{X}_t^{\mathbf{w}}\}_{t=1}^T$, где $\mathbf{X}_t^{\mathbf{w}} = \mathbf{X}_t \odot \mathbf{w}$, а \odot – покомпонентное произведение векторов Адамара.
2. Подсчитываем матрицу $\mathbf{P} = \mathbf{P}(\mathbf{w}) \in \mathbb{R}^{n \times N}$, строки которой равняются компонентам собственных векторов матрицы $\mathbf{C}^{\mathbf{w}}$, соответствующих первым n максимальным собственным значениям.
3. Элементы $a_{i,j}$ матрицы $\mathbf{A} = \mathbf{A}(\mathbf{w})$ равны $a_{i,j} = p_{i,j} w_j, i = 1, \dots, n, j = 1, \dots, N$, где $p_{i,j}$ – элементы матрицы \mathbf{P} .
4. Элементы $b_{i,j}$ матрицы $\mathbf{B} = \mathbf{B}(\mathbf{w})$ равны $b_{i,j} = p_{j,i} \frac{1}{w_i}, i = 1, \dots, n, j = 1, \dots, N$.

5. “Восстановленные” вектора $\widehat{\mathbf{X}}_t = \mathbf{B}\mathbf{A}\mathbf{X}_t$, $t = 1, \dots, T$.

В некоторых приложениях оказывается важным, чтобы матрицы $\mathbf{A} = \mathbf{A}(\mathbf{w})$ и $\mathbf{B} = \mathbf{B}(\mathbf{w})$ определялись не как решение задачи минимизации (5), а как решение задачи минимизации

$$\delta_1(n, \rho_\infty, \mathbf{w}) \rightarrow \min_{\mathbf{A}: \mathbf{A}\mathbf{B}=\mathbf{I}} \quad (6)$$

или задачи минимизации

$$\delta_2(n, \rho_\infty, \mathbf{w}) \rightarrow \min_{\mathbf{A}: \mathbf{A}\mathbf{B}=\mathbf{I}}. \quad (7)$$

Однако, явное решение для (6) или (7) получить сложно.

Таким образом, задача состоит в том, чтобы используя матрицы $\mathbf{A} = \mathbf{A}(\mathbf{w})$ и $\mathbf{B} = \mathbf{B}(\mathbf{w})$, являющиеся решением задачи минимизации (5) для произвольного вектора весов \mathbf{w} (т.е. моделируя многообразие \mathcal{M} с помощью взвешенных главных компонент), уменьшить ошибки $\delta_1(n, \rho_\infty, \mathbf{w}^*)$ и $\delta_2(n, \rho_\infty, \mathbf{w}^*)$ (для какого-то конкретного вектора весов \mathbf{w}^* , задаваемого с учетом специфики данных) за счет специального подбора \mathbf{w} в зависимости от \mathbf{w}^* .

4. ИТЕРАТИВНЫЙ АЛГОРИТМ ПОДБОРА ВЕСОВ

Идея подбора весов \mathbf{w} заключается в том, чтобы итеративно изменять веса w_j , $j = 1, \dots, N$ пропорционально $e_j^w = w_j^* \cdot \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \widehat{X}_{j,t}|$ (взвешенная усредненная по выборке ошибка восстановления по j -ой координате). Очевидно, что веса w_j для тех индексов j , для которых ошибка e_j^w большая, будут расти быстрее относительно других весов. Как следствие, по соответствующим координатам ошибка восстановления будет уменьшаться, а по другим – увеличиваться. Значит, максимальная ошибка $\rho_\infty(\mathbf{X}_t, \widehat{\mathbf{X}}_t, \mathbf{w}^*) = \max_{j=1, \dots, N} \left\{ w_j^* |X_{j,t} - \widehat{X}_{j,t}| \right\}$ в среднем по выборке будет уменьшаться, что приведет к уменьшению ошибок $\delta_1(n, \rho_\infty, \mathbf{w}^*)$ и $\delta_2(n, \rho_\infty, \mathbf{w}^*)$.

Алгоритм

Вход:

- a) выборка $\{\mathbf{X}_t\}_{t=1}^T$
- b) вектор весов \mathbf{w}^*
- c) количество ГК n
- d) максимальное количество итераций K
- e) максимально допустимое значение веса w^{\max}
- f) функционал $\delta \in \{\delta_1, \delta_2\}$, измеряющий ошибку восстановления (см. (3), (4))

Выход:

- a) оптимальный вектор весов \mathbf{w}^{opt}
- b) матрицы $\mathbf{A}(\mathbf{w}^{\text{opt}})$ и $\mathbf{B}(\mathbf{w}^{\text{opt}})$ (см. (5))

1. $\mathbf{w}^{(0)} \leftarrow (1, \dots, 1)$
2. $k \leftarrow 0$
3. **repeat**
 - a) подсчитать $\mathbf{A} = \mathbf{A}(\mathbf{w}^{(k)})$ и $\mathbf{B} = \mathbf{B}(\mathbf{w}^{(k)})$ (см. (5))
 - b) $\widehat{\mathbf{X}}_t \leftarrow \mathbf{B}\mathbf{A}\mathbf{X}_t$, $t = 1, \dots, T$
 - c) $E^{(k)} \leftarrow \delta(n, \rho_\infty, \mathbf{w}^*)$

- d) $e_j^w \leftarrow w_j^* \cdot \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \widehat{X}_{j,t}|, j = 1, \dots, N$
 - e) $w_j^{(k+1)} \leftarrow w_j^{(k)} \cdot \left(1 + \frac{e_j^w}{\sum_{i=1}^N e_i^w} \right), j = 1, \dots, N$
 - f) $k \leftarrow k + 1$
 - g) $w_j^{(k)} \leftarrow \frac{w_j^{(k)}}{\min_{i=1, \dots, N} w_i^{(k)}}, j = 1, \dots, N$
4. **until**
- a) $k < K$ и
 - b) $\forall j \in \{1, \dots, N\} \rightarrow w_j^{(k)} < w^{\max}$.
5. $k_{\text{opt}} \leftarrow \arg \min_{i \in \{0, \dots, k-1\}} E^{(i)}$
6. $\mathbf{w}^{\text{opt}} \leftarrow \mathbf{w}^{(k_{\text{opt}})}$
7. $\mathbf{A} \leftarrow \mathbf{A}(\mathbf{w}^{\text{opt}}), \mathbf{B} \leftarrow \mathbf{B}(\mathbf{w}^{\text{opt}})$

Условие остановки, приведенное в пункте 4.b алгоритма, необходимо потому, что в результате нормировки компонент вектора весов согласно формуле пункта 3.i алгоритма значения некоторых весов могут существенно увеличиться, что неизбежно приведет к неустойчивости вычислений. Величину параметра w^{\max} следует выбирать исходя из типичного масштаба выборочных данных.

5. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Проиллюстрируем эффективность предложенного алгоритма на примере сжатия выборки, состоящей из геометрических описаний профилей крыла самолета. Сжатие геометрического описания самолета необходимо для применения когнитивной технологии быстрых расчетов [20, 21], позволившей на основе данных построить суррогатную модель самолета [22, 23, 24, 25], с помощью которой можно определять аэродинамические характеристики самолета для различных вариантов его компоновки и в различных условиях функционирования с тем, чтобы проводить сравнение различных технических решений касательно компоновки самолета.

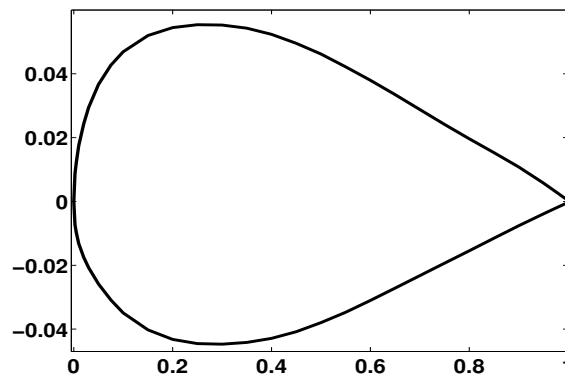


Рис. 1. График профиля BOEING 737 MIDSPAN.

Геометрическое описание профиля крыла самолета задается вектором значений ординат верхней и нижней частей контура профиля (в рассматриваемом случае количество ординат равнялось 57) при фиксированных значениях абсцисс.

Отметим, что в силу аэродинамических соображений [26, 27] для измерения ошибки восстановления должен быть использован именно функционал $\delta_1(n, \rho_\infty, \mathbf{w}^*)$ (или $\delta_2(n, \rho_\infty, \mathbf{w}^*)$).

При этом значения весов w_j^* , соответствующие ординатам точек, расположенных в передней и задней частях профиля, должны быть больше значений остальных весов, поскольку именно эти элементы профиля (передняя и задняя части) должны восстанавливаться с большей точностью.

При расчетах использовалась выборка профилей объема 1368 (данные были получены из открытых источников [28, 29], см. также ссылки на литературу в [29]). На рис. 1 изображен график типичного профиля из выборки.

Использовались следующие значения параметров алгоритма:

- а) веса $w_j^* = 5$, $j = 1, \dots, 7, 50, \dots, 57$, $w_j^* = 2$, $j = 25, \dots, 35$ и $w_j^* = 1$ при остальных j
- б) количество ГК $n = 6$
- в) максимальное количество итераций $K = 500$
- г) максимально допустимое значение веса $w^{\max} = 10^4$
- д) функционал δ_1 (см. (3))

Время выполнения (основные характеристики компьютера – Intel Core Duo Processor T7200, 2.00 GHz, 2×1024 МВ/Мо, 667 МГц) в системе MatLab 7.4.0 (R2007a) составило 23.2156 сек.

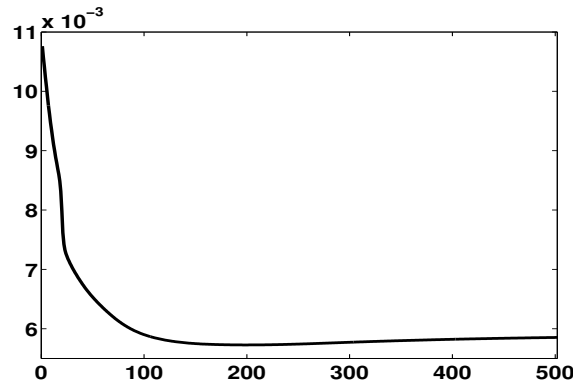


Рис. 2. График значений ошибки $E^{(i)}$, $i = 0, \dots, K - 1$.

На рис. 2 изображен график значений ошибки $E^{(i)}$, $i = 0, \dots, K - 1$. Видно, что сначала ошибка падает (примерно в 2 раза), а потом начинает расти, при этом $k_{\text{opt}} = 196$. На рис. 3 изображены значения w_j^{opt} , $j = 1, \dots, 57$. Видно, что значения весов, номера которых соответствуют точкам передней и задней частей профиля, имеют большие значения по сравнению с остальными весами. На рис. 4 изображены значения ошибок $e_j^w = w_j^* \cdot \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \hat{X}_{j,t}|$, $j = 1, \dots, 57$ при $\mathbf{w} = \mathbf{w}^{(0)}$ (штрих-пунктирная линия) и $\mathbf{w} = \mathbf{w}^{\text{opt}}$ (сплошная линия). На рис. 5 изображены значения ошибок $e_j = \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \hat{X}_{j,t}|$, $j = 1, \dots, 57$ при $\mathbf{w} = \mathbf{w}^{(0)}$ (штрих-пунктирная линия) и $\mathbf{w} = \mathbf{w}^{\text{opt}}$ (сплошная линия). Видно, что при применении алгоритма средняя (по выборке) взвешенная максимальная (по координате) ошибка действительно падает.

6. ЗАКЛЮЧЕНИЕ

Итак, в статье разработан итеративный алгоритм подсчета взвешенных главных компонент, который позволяет существенно уменьшить среднюю (по выборке) взвешенную максимальную (по координате) ошибку восстановления векторов выборки по сжатым данным. Приведенные результаты моделирования на примере сжатия геометрического описания выборки профилей крыла самолета показали эффективность предложенного алгоритма.

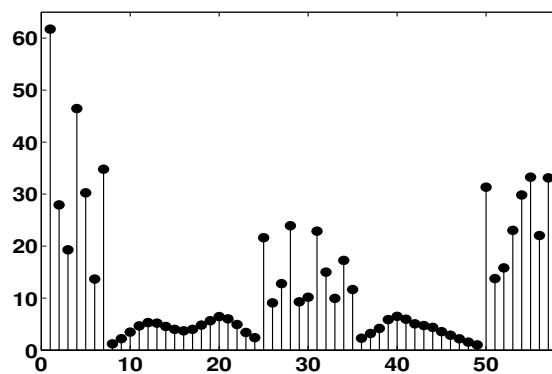


Рис. 3. График значений $w_j^{\text{opt}}, j = 1, \dots, 57$.

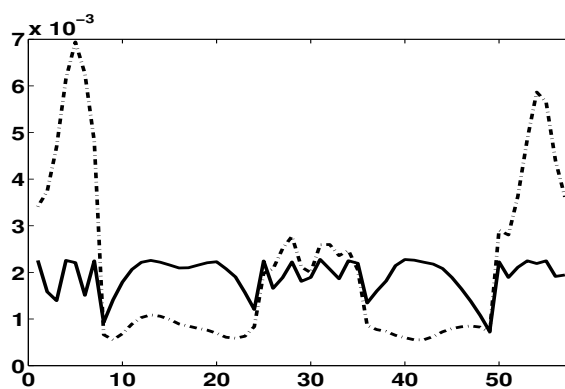


Рис. 4. График значений ошибок $e_j^w = w_j^* \cdot \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \hat{X}_{j,t}|, j = 1, \dots, 57$ при $\mathbf{w} = \mathbf{w}^{(0)}$ (штрих-пунктирная линия) и $\mathbf{w} = \mathbf{w}^{\text{opt}}$ (сплошная линия).

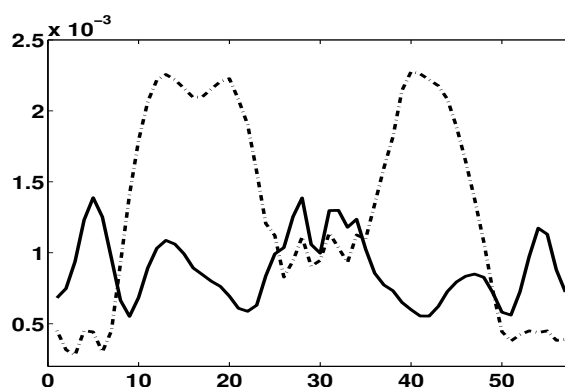


Рис. 5. График значений ошибок $e_j = \frac{1}{T} \sum_{t=1}^T |X_{j,t} - \hat{X}_{j,t}|, j = 1, \dots, 57$ при $\mathbf{w} = \mathbf{w}^{(0)}$ (штрих-пунктирная линия) и $\mathbf{w} = \mathbf{w}^{\text{opt}}$ (сплошная линия).

В дальнейшем необходимо:

1. Найти точное решение задач оптимизации (6) и/или (7)
2. Попытаться обобщить предложенный алгоритм на случай нелинейных ГК (см. [30, 31, 32])

СПИСОК ЛИТЕРАТУРЫ

1. Fukunaga, K., *Introduction to Statistical Pattern Recognition*, San Diego: Academic, 1990.
2. Jimenez, L.O. and Landgrebe, D.A., Supervised Classification in High-dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate data. *IEEE Transactions on Systems, Man and Cybernetics*, 1997, no. 28(1), pp. 39–54.
3. van der Maaten, L.J.P., Postma, E.O., and van den Herik, H.J., Dimensionality Reduction: A Comparative Review, submitted to *Neurocognition*, 2008 (http://www.cs.unimaas.nl/l.vandermaaten/Laurens_van_der_Maaten/Publications.html).
4. Fodor, I.K., *A Survey of Dimension Reduction Techniques*, LLNL Technical Report, 2002, UCRL-ID-148494.
5. Bellman, R., *Adaptive Control Processes: A Guided Tour*, Princeton: Princeton Univ. Press, 1961.
6. Scott, D.W. and Thompson, J.R., Probability Density Estimation in Higher Dimensions, in *Computer Science and Statistics: Proceedings of the Fifteenth Symposium on the Interface*, Gentle, J.E., Ed., New York: Elsevier, 1983, pp. 173–179.
7. Scott, D.W., *Multivariate Density Estimation. Theory, Practice, and Visualization*, New York: Wiley, 1992.
8. Silverman, B.W., *Density Estimation for Statistics and Data Analysis*, New York: Chapman & Hall, 1986.
9. Jackson, J.E., *A User's Guide to Principal Components*, New York: Wiley, 1991.
10. Jolliffe, I.T., *Principal Component Analysis, Springer Series in Statistics*, Berlin: Springer-Verlag, 1986.
11. Айвазян, С.А., Бухштабер, В.М., Енюков, И.С., Мешалкин, Л.Д., *Прикладная статистика: классификация и снижение размерности*, М.: Финансы и статистика, 1989.
12. Collins, M., Dasgupta, S., and Schapire, R.E., A Generalization of Principal Components Analysis to the Exponential Family, in *Advances in Neural Information Processing Systems 14 (NIPS)*, Dietterich, T.G., Becker, S., and Zoubin Ghahramani, Eds., Cambridge: MIT Press, 2002.
13. Sajama and Orlitsky, A., Semi-parametric Exponential Family PCA. Advances, in *Neural Information Processing Systems 17 (NIPS)*, Vancouver, 2004
14. Skocaj, D., Leonardis, A., and Bischof, H., Weighted and Robust Learning of Subspace Representations, *Pattern Recognition*, 2007, vol. 40, issue 5, pp. 1556–1569.
15. Roweis, S., EM Algorithms for PCA and SPCA, *Proc. 1997 Conf. on Advances in Neural Information Processing Systems 10*, Denver, 1998, pp. 626–632.
16. Zou, H., Hastie, T., and Tibshirani, R., Sparse Principal Component Analysis, *J. Computational and Graphical Statistics*, 2006, vol.15, no. 2, pp. 265–286.
17. Chipman, H.A. and Gu, H., Interpretable Dimension Reduction, *J. Appl. Statistics*, 2005, vol. 32, no. 9, pp. 969–987.
18. De la Torre, F. and Black, M. J., Robust Principal Component Analysis for Computer Vision, in *Int. Conf. on Computer Vision, ICCV-2001*, Vancouver, 2001, pp. 362–369.
19. Skocaj, D. and Leonardis, A., Weighted and Robust Incremental Method for Subspace Learning, *9th IEEE International Conference on Computer Vision (ICCV 2003)*, Nice, France, 2003, pp. 1494–1501.
20. Кулешов, А.П., Когнитивные технологии в основанных на данных адаптивных моделях сложных объектов, *Информационные технологии и вычислительные системы*, 2008, вып. 1.

21. Кулешов, А.П., Технология быстрого вычисления характеристик сложных технических объектов, *Информационные технологии*, 2006. № 3, стр. 4–11.
22. Вышинский, В.В., Свириденко, Ю.Н., Применение технологии быстрого вычисления характеристик сложных технических объектов для расчета аэродинамических характеристик самолета, *приложение к журналу "Информационные технологии"*, 2006, Вып. 3, стр. 12–17.
23. Бернштейн, А.В., Вышинский, В.В., Кулешов, А.П., Свириденко, Ю.Н., Применение искусственных нейронных сетей для определения нагрузок по крылу пассажирского самолета на режиме крейсерского полета, *Труды ЦАГИ им. проф. Н.Е. Жуковского, "Применение искусственных нейронных сетей в задачах прикладной аэродинамики"*, Москва, 2008, Выпуск № 2678.
24. Бернштейн, А.В., Вышинский, В.В., Кулешов, А.П., Свириденко, Ю.Н., Быстрый метод аэродинамического расчета для задач проектирования, *Труды ЦАГИ им. проф. Н.Е. Жуковского, "Применение искусственных нейронных сетей в задачах прикладной аэродинамики"*, Москва, 2008, Выпуск № 2678.
25. Свириденко, Ю.Н., Применение сжатия данных для случайной генерации объектов с заданными аэродинамическими характеристиками, *Труды ЦАГИ им. проф. Н.Е. Жуковского, "Применение искусственных нейронных сетей в задачах прикладной аэродинамики"*, Москва, 2008, Выпуск № 2678.
26. Кашин Г. М., Пшеничнов Г. И., Флеров Ю. А. *Методы автоматизированного проектирования самолета*. М.: Машиностроение, 1979.
27. Давыдов Ю.В., Злыгарев В.А. *Геометрия крыла*. М.: Машиностроение, 1987.
28. Кашафутдинов С.Т., Лушин В.Н. *Атлас аэродинамических характеристик крыловых профилей*. Новосибирск: СО РАСХН, 1994.
29. Mason W. H. *Applied computational aerodynamics*. Lecture Notes. Virginia Polytechnic Institute and State University. Blacksburg, 1995 (http://www.aoe.vt.edu/~mason/Mason_f/CAtxtTop.html).
30. Kramer, M.A., Nonlinear Principal Component Analysis Using Autoassociative Neural Networks, *AIChE J.*, 1991, no. 37(2), pp. 233–243.
31. Schölkopf, B., Smola, A., and Müller, K.-R., Nonlinear Component Analysis as a Kernel Eigenvalue Problem, *Neural Computation*, 1998, no. 10(5), pp. 1299–1319.
32. DeMers, D. and Cottrell, G.W., Non-linear Dimensionality Reduction, in *Advances in Neural Information Processing Systems*, Hanson, S.J., Cowan, J.D., and Giles, C.L., Eds., San Mateo: Morgan Kaufmann, 1993, vol. 5, pp. 580–587.